Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 5443–5448, May 1996
Biochemistry

Conserved motifs in prokaryotic and eukaryotic polypeptide release factors: tRNA–protein mimicry hypothesis

(translation termination/stop codon recognition/elongation factor G/molecular mimicry)

KOICHI ITO, KANAE EBIHARA, MAKIKO UNO, AND YOSHIKAZU NAKAMURA*

Department of Tumor Biology, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108, Japan

Communicated by Charles Yanofsky, Stanford University, Stanford, CA, January 31, 1996 (received for review November 1, 1995)

ABSTRACT **Translation termination requires two codon-specific polypeptide release factors in prokaryotes and one omnipotent factor in eukaryotes. Sequences of 17 different polypeptide release factors from prokaryotes and eukaryotes were compared. The prokaryotic release factors share residues split into seven motifs. Conservation of many discrete, perhaps critical, amino acids is observed in eukaryotic release factors, as well as in the C-terminal portion of elongation factor (EF) G. Given that the C-terminal domains of EF-G interacts with ribosomes by mimicry of a tRNA structure, the pattern of conservation of residues in release factors may reflect requirements for a tRNA-mimicry for binding to the A site of the ribosome. This mimicry would explain why release factors recognize stop codons and suggests that all prokaryotic and eukaryotic release factors evolved from the progenitor of EF-G.**

The termination of translation in bacteria requires two codon-specific peptide chain release factors (RF): release factor-1 (RF-1; UAG/UAA specific) and release factor-2 (RF-2; UGA/UAA specific) (1). A third factor, release factor-3 (RF-3), is known to stimulate the activities of RF-1 and RF-2 and to bind guanine nucleotides, but is not codon-specific (2–6). The mechanism of stop codon recognition by RFs is unknown and is of considerable interest since it entails protein–RNA recognition rather than the well-understood mRNA–tRNA interaction in codon-anticodon pairing (7, 8).

The existence of a protein with RF activity in eukaryotes was demonstrated some 20 years ago in rabbit reticulocytes (9). After two decades of investigation, a eukaryotic protein family with the properties of RFs, designated eRF-1, has been discovered (10). Two members of the eRF-1 family, one from humans and the other from *Xenopus laevis*, have been purified and shown to catalyze polypeptide release at all three stop codons *in vitro*. Another protein in this family, Sup45 of the budding yeast *Saccharomyces cerevisiae*, is involved in omnipotent suppression during translation (for a review, see ref. 11). However, the discoverers of the eRF-1 family failed to detect any sequence resemblance between bacterial RFs (including a yeast mitochondrial RF, mRF-1) and eukaryotic RFs, and speculated that the proteins responsible for eukaryotic termination evolved independently from the RFs of prokaryotes (10).

It appeared unlikely that eRF-1 is the only RF required for polypeptide termination in eukaryotes because no member of the eRF-1 family contained a consensus G-domain structure. A GTP requirement for polypeptide release was demonstrated some 20 years ago in rabbit reticulocytes (9). Eukaryotic counterparts of bacterial RF-3 have been proposed recently and referred to as eRF-3 (12). The reported eRF-3 family includes a Sup35 protein of *S. cerevisiae,* which carries G-domain motifs and is involved in omnipotent suppression of nonsense codons (for a review, see ref. 11).

Can the current computer programs used for sequence comparison, as designed, predict conserved amino acids at discrete positions in comparisons of multiple random sequences? It seems unlikely to us that the currently used computer programs would recognize single conserved amino acids when the number and diversity of protein sequences is increased, because the algorithms used are essentially based on one-to-one comparison of letters or words of finite length. Here, we approach this problem by identifying first "by computer" the conserved amino acids in prokaryotic RFs, and then asking "by eye" whether these residues are also present in eukaryotic RFs. This approach provided us with clues that lead to universally conserved motifs in RFs, part of which may reflect requirements for molecular mimicry of a tRNA structure.

MATERIALS AND METHODS

***Sup45* Gene of *Schizosaccharomyces pombe*.** The *Sup45* gene of the fission yeast *Sch. pombe* was cloned by DNA hybridization using the *S. cerevisiae SUP45* probe. The size of the *Eco*RI fragment encoding *S. pombe* Sup45 was estimated to be ≈10 kb by Southern blot hybridization using the *S. cerevisiae SUP45* probe isolated from pUKC802 (gift from M. F. Tuite, University of Kent; see ref. 13). A shotgun library of *Sch. pombe Eco*RI DNAs in λDASH II vector (Stratagene) was plaque-screened. After several rounds of screening, positive phage clones were characterized by genetic and DNA sequencing analyses. The 6-kb *Sal*I–*Sal*I fragment subcloned in plasmid pRS316 (14) encoded a Sup45 homolog which is able to complement intergenerically the *S. cerevisiae sup45* mutation.

***Streptomycetes* RF-2 Gene.** Streptomycetes are mycelial, Gram-positive bacteria that produce a wide variety of antibiotics and bioactive compounds. The RF-2 gene of *Streptomyces coelicolor* A3(2) was found in a 6.3-kb DNA fragment also encoding protein kinases (15). DNA sequence analysis revealed that *Streptomyces* RF-2 comprises 368 amino acids and is 43–45% identical to *Escherichia coli*, *Salmonella typhimurium*, or *Bacillus subtilis* RF-2. The *Streptomyces* RF-2 gene was able to complement the *E. coli prfB* mutations (15).

Dominant Lethal Mutations in *E. coli* RF-2 Gene. Plasmid pSUIQ-RF2+ is an isopropyl β-D-thiogalactoside (IPTG)-controllable *prfB*$^+$ expression plasmid that contains the entire *prfB*$^+$ encoding sequence of *E. coli* under the *lac* promoter and the *lacI*^{q} sequence. The pSUIQ-RF2+ complements the *E. coli*

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RF, release factor; RF-1, release factor 1; RF-2, release factor 2; RF-3, release factor 3; eRF-1, eukaryotic release factor 1; eRF-3, eukaryotic release factor 3; mRF-1, mitochondrial release factor 1; IPTG, isopropyl β-D-thiogalactoside; EF, elongation factor.
Data deposition: The sequences reported in this paper have been deposited in the GenBank data base (accession nos. D63883 and D26540).

*To whom reprint requests should be addressed. e-mail: nak@hgc.ims.u-tokyo.ac.jp.

prfB mutations in the presence of IPTG. The derivatives carrying the *prfB2070* and *prfB2130* mutations were constructed by creating single amino acid substitutions in pSUIQ-RF2+ DNA at positions 207 and 213 by site-directed mutagenesis (16) by using primers 5′-AAAGCCCGACT-GACTCCG-3′ and 5′-GCGGTCGTATCCACACGT-3′, respectively (base substitutions are underlined). Details of these and other plasmid constructions will be published elsewhere.

RESULTS AND DISCUSSION

***S. pombe Sup45* Gene Structure.** Two novel RF genes that are from species distantly related to those with known RF genes were cloned and sequenced to facilitate the sequence comparisons. One encoded the Sup45 homolog of the fission yeast, *Sch. pombe*, and the other encoded an RF-2 of *Streptomyces coelicolor* (15). The *sup45* gene of *Sch. pombe* encodes a polypeptide of 434 amino acids and is homologous to other eRF-1 genes. It shares 63% amino acid identity to *S. cerevisiae* Sup45 and complements the *S. cerevisiae sup45* mutation. The *S. cerevisiae* strain MT557/1d (*sal4-2 ura3-1 ade2-1 leu2-3, 112 MATa*) cannot grow at 37°C owing to a temperature-sensitive allele, *sal4-2*, in the *SUP45* gene (17). MT557/1d cells transformed at 25°C with a pRS316 derivative carrying the cloned Sup45 homolog of *Sch. pombe* became viable at 37°C, showing intergeneric complementation of *S. cerevisiae sup45* with the *Sch. pombe* counterpart.

Conserved Motifs in Prokaryotic and Eukaryotic RFs. Eleven prokaryotic RF sequences (18–23) and *S. cerevisiae* mRF-1 (24), were compared by the BESTFIT or PILEUP programs from the Genetics Computer Group program package (25). The similarity score plot showed that there are at least seven highly conserved regions in the prokaryotic and lower eukaryotic RFs (Fig. 1). These results lead us to propose a seven-domain model for prokaryotic RF structure.

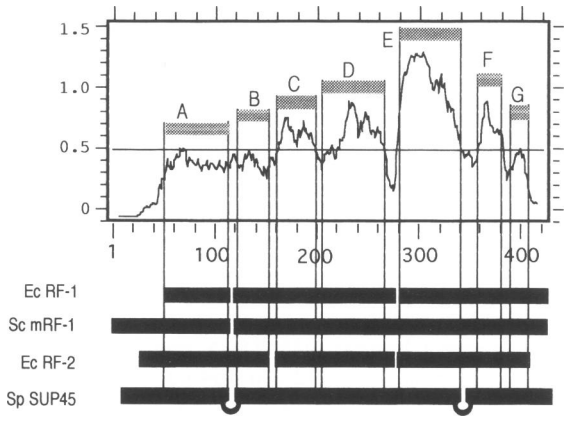

FIG. 1. The average similarity plot of the 11 prokaryotic RF sequences and *S. cerevisiae* mRF-1: *E. coli* RF-1 (Ec RF-1) and RF-2 (Ec RF-2), *S. typhimurium* RF-1 and RF-2, *B. subtilis* RF-2, *Streptomyces* RF-2 (GenBank accession no. D26540), *Mycobacterium leprae* RF-1 (GenBank accession no. U01800), *Mycoplasma genitalium* RF-1 (The Institute for Genomic Research Database accession no. MG258), *Haemophilus influenzae* RF-1 (GenBank accession no. L46192), and *Haemophilus influenzae* RF-2 (GenBank accession no. L45847), *E. coli* RF-H (an RF homolog, see ref. 24), and *S. cerevisiae* mRF-1 (Sc mRF-1). Seven conserved regions are assigned as described in the text and are indicated by gray boxes. Structural alignments of representative RFs are shown by closed boxes, with gaps given for optimal sequence similarity to eRF-1 as described in Fig. 2. The average similarity score along the entire sequence is provided by the horizontal line. Comparison scores (expressed in standard deviations) were calculated using the PILEUP program from the Genetics Computer Group program package (25). Sp SUP45 represents *Sch. pombe* Sup45.

Five eRF-1 sequences from human, *Xenopus laevis*, *Arabidoposis thaliana*, *S. cerevisiae* (10), and *Sch. pombe* (this study) were aligned first by using the BESTFIT or PILEUP programs. Then, we compared "by eye" the yeast Sup45 sequences with the prokaryotic alignment by assuming appropriate, but minimal, gaps to achieve an optimal match. Residues that were identical or at which there were conservative replacements in comparisons with the yeast Sup45 sequences are shown in Fig. 2. Seven prokaryotic RFs and four eRF-1s that are distantly related are compared. Most of the amino acids from the eukaryote proteins assigned using this strategy coincided with those highly conserved among the prokaryotic RF sequences, suggesting that there is a pattern of conservation of many discrete, perhaps critical, amino acids in all RFs. One third of the amino acid positions of RFs were assigned to those having identical residues in both prokaryotic and eukaryotic RFs, and half of them were assigned to those having similar residues. In contrast, amino acids located at positions between these conserved residues are divergent between prokaryotic RFs and eRF-1s.

It is perhaps highly relevant that the four gaps identified in comparisons with Sup45 (eRF-1) exactly coincide with the junctions between conserved prokaryotic regions, A-B, B-C, D-E, and E-F (see Fig. 2). The gap between regions D-E corresponds to a papain-sensitive cleavage site in bacterial RFs (27), perhaps reflecting the boundary of predicted domains in protein conformation. Moreover, all 14 of the RF mutations whose amino acid substitutions have been determined (17, 28–33) affect the conservative amino acids or neighboring residues (Table 1; positions marked by asterisks in Fig. 2). These results provide evolutionary evidence for the biological and structural significance of these sequence comparisons of RFs.

One example of the conservative amino acids is Arg^{137} of *E. coli* RF-1 (position 215 in the given coordinate of Fig. 2; region C). This residue is altered by the temperature-sensitive *prfA1* mutation of *E. coli*. Zhang *et al.* (33) have shown that the *rplL564* allele, which codes for an altered ribosomal protein L7/L12, can suppress *prfA1*, while other *rplL* mutant alleles do not. These data may suggest the role of this conserved residue, Arg^{137}, for ribosome-binding during translation termination.

Obviously, the acquisition of new RF sequences, such as those of *Streptomyces* RF-2 and *Sch. pombe* Sup45, facilitate the comparison, but the strategy of performing multiple sequence comparisons by computer and by eye goes beyond computer capabilities because of limitations of the processing speed and commonly available algorithms. We strongly suggest that there is sequence conservation in the eRF-1 prokaryotic RF family, indicating the possibility of a mechanistic resemblance between the eukaryotic and prokaryotic processes, and that the members of the family evolved from a common ancestor.

tRNA–Protein Mimicry. The same strategy for sequence comparison was applied to prokaryotic RFs and EF-G because these proteins might be expected to share structural features necessary for binding to the A site of the ribosome. Here, we compared EF-G with the prokaryotic RF alignment by postulating minimal gaps for an optimal match. The data revealed the conservation pattern of many discrete amino acids in RFs and the C-terminal part of EF-G (Fig. 3*A*). Half of these conservative amino acids were also highly conserved in prokaryotic EF-Gs (positions marked by asterisks in Fig. 3*A*). Ævarsson *et al.* (34) and Czworkowski *et al.* (35) have solved the three-dimensional structure of *Thermus thermophilus* EF-G and proposed six subdomains, G, G′, and II–V (Fig. 4). The RF regions predicted to be conserved in this paper are located in domains III–V. The following observations support the evolutionary validity of this alignment. First, the predicted gap of 15 amino acids coincides with the junction between domains III–IV of EF-G (see Fig. 3*A*). Second, the boundaries

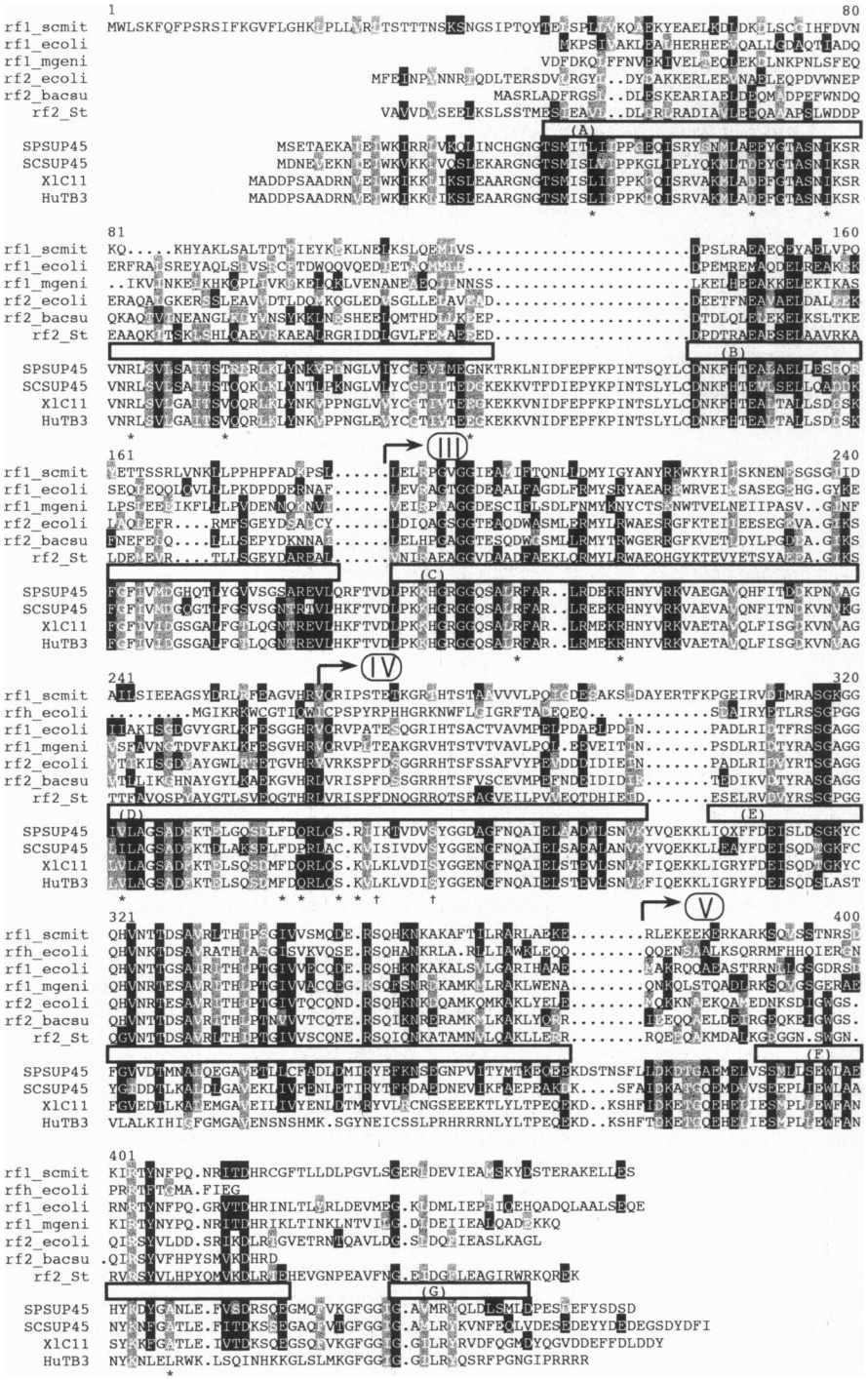

FIG. 2. Comparison of the amino acid sequences of prokaryotic and eukaryotic RFs. The similarity alignments of prokaryotic RFs and eRF-1s were separately accomplished by using the BESTFIT or PILEUP programs from the Genetics Computer Group program package (25), and then the eRF-1 sequence, as represented by yeast Sup45 sequences, was compared "by eye" with the prokaryotic RF alignment by inserting appropriate gaps (as indicated by dots). Any divergent residues at the same positions in the two yeast Sup45 sequences are permitted for the assignment since *Sch. pombe* Sup45 complements the Sup45 mutation of *S. cerevisiae*. Identical residues are boxed in black. Conserved residues (26), grouped as I-L-M-V, D-E, R-K, S-T, F-Y and G-A, are boxed in gray. The number of the amino acid position is counted from the N-terminal Met of the mRF-1 sequence. Asterisks indicate 14 RF mutant alleles sequenced thus far, and daggers indicate positions of 2 dominant lethal mutations in *E. coli prfB* isolated in this study (see Table 1). rf1_scmit, *S. cerevisiae* mitochondrial mRF-1; rf1_ecoli, *E. coli* RF-1; rf2_ecoli, *E. coli* RF-2; rf2_bacsu, *B. subtilis* RF-2; rfh_ecoli, *E. coli* RF-H; rf2_St, *Streptomyces* RF-2; rf1_mgeni, *M. genitalium* RF-1; SPSUP45, *Sch. pombe* Sup45; SCSUP45, *S. cerevisiae* Sup45; XlC11, *X. laevis* eRF-1; HuTB3, human eRF-1. The conserved regions A–G are shown by boxes under the prokaryotic RF alignment. Structural boundaries predicted by sequence comparison with elongation factor (EF) G domains III-V are shown by arrows.

of domains II–III and IV–V assigned by computer match the two sequence gaps between conserved RF regions B–C and E–F, respectively (see Fig. 2).

The three-dimensional structure of the ternary complex of Phe-tRNA, *Thermus aquaticus* EF-Tu, and the nonhydrolyzable GTP analog, GDPNP, has been described recently (36). This structure was almost completely superimposable with EF-G (36), showing that domains III, IV, and V appear to mimic the shapes of the acceptor stem, anticodon helix, and T stem of tRNA in the ternary complex, respectively. Domain IV of EF-G forms a protruding 'rod' conformation (34, 35), which is similar to the shape of the anticodon arm of tRNA (Fig. 4). When the Phe-tRNA structure was aligned with the C-terminal part of EF-G, using the C^{α} coordinates from domains III–V, the two structures were superimposable except for minor differences. This resemblance of part of EF-G and tRNA represents an instance of structural homology between a protein and a nucleic acid. Functional mimicry of a major autoantigenic epitope of the human insulin receptor by RNA has been described (37), and protein mimicry of DNA has been shown in the crystal structure of the uracil-DNA glycosylase-uracil glycosylase inhibitor protein complex (38, 39). These independent findings suggest a novel concept of "molecular mimicry between nucleic acid and protein."

These observations prompted us to propose that the RF region that has the conservation pattern common to domain IV may be a tRNA-mimicry domain necessary for RF binding to the A site of the ribosome. Several lines of evidence favor this proposal. (*i*) tRNA:EF–Tu:GTP, EF–G:GTP complexes, and RF bind to the A site of ribosomes during translation (40).

Table 1. RF mutations determined by the sequence analysis

Domain	RF*	Coordinate	Allele (ref.)	Amino acid substitution	Mutant property†	Counterpart domain in EF-G‡
A	Sc Sup45	52	*sup1-36* (28)	Leu^{34}→Ser	Nonsense suppressor, Ts	
A	Sc Sup45	69	*asu255* (28)	Asp^{51}→Asn	Suppressor of *sup1-36*	
A	Sc Sup45	77	*asu198-1* (28)	Ile^{59}→Thr	Suppressor of *sup1-36*	
A	Sc Sup45	83	*sup45-3* (26)	Arg^{65}→Cys	Nonsense suppressor, respiratory deficiency	
A	Sc Sup45	93	*asu198-2* (28)	Thr^{75}→Ile	Suppressor of *sup1-36*	
A	Ec RF-2	119	*prfB1* (25)	Glu^{89}→Lys	UGA suppressor	
C	Ec RF-2	204	*prfB3* (25)	Asp^{143}→Asn	UGA suppressor	III
C	Ec RF-1	215	*prfA1* (30)	Arg^{137}→Pro	UAG/UAA suppressor, Ts	III
D	Sc Sup45	242	*sal4-2* (16)	Ile^{222}→Ser	Nonsense suppressor, Ts	III
D	St RF-1	259	*prfA10* (27)	Gly^{180}→Ser	UAG suppressor, Ts	III/IV
D	St RF-1	261	*prfA8* (27)	His^{182}→Tyr	UAG suppressor	III/IV
D	Sc mRF-1	265	*mrf1-1* (29)	Arg^{231}→Lys	Nonsense suppressor, respiratory deficiency	III/IV
D	Sc mRF-1	267	*mrf1-2* (29)	Pro^{233}→Leu	Nonsense suppressor, respiratory deficiency	III/IV
D	Ec RF-2	269	*prfB2070*	Phe^{207}→Thr	Dominant lethal	IV
D	Ec RF-2	275	*prfB2130*	Arg^{213}→Ile	Dominant lethal	IV
F	Ec RF-2	407	*prfB286* (25)	Leu^{328}→Phe	UGA suppressor, Ts	V

RF mutations examined by DNA sequence analysis are summarized according to the new coordinate and by the domain classification as shown in Fig. 2.

*Sc, *S. cerevisiae*; Ec, *E. coli*; St, *Sal. typhimurium*.

†Ts, temperature-sensitive growth.

‡Structural domains III–V of EF-G. III/IV indicates the boundary between domains III and IV.

(*ii*) Hydrophilicity plots of domain IV of EF-G and its RF counterparts are homologous (Fig. 3*B*), suggesting that these regions have a similar distribution of surface charge. (*iii*) Predicted secondary structures of these RF regions show significant similarity, consisting of two tandem repeats of the β strand-α helix motif (data not shown). This feature might relate to some extent to the three-dimensional structure of domain IV of EF-G that has an unusual βαβ topology, with a left-handed crossover connection of two central β-strands (34, 35). (*iv*) *Sch. pombe* Sup45 also shows similar features in hydrophilicity and secondary structure predictions (see Fig. 3*B*; data not shown).

Anticodon Mimicry. It is tempting to speculate further that several structural features of these proteins are required for mimicry of tRNA. We would propose tRNA-mimicry determinants for these crucial elements using the analogy to tRNA-identity determinants. The most critical of these mimicry determinants would be the anticodon-mimicry element in the protein, which probably consists of some of the conserved or less conserved residues in the RF counterpart of domain IV.

Ten conserved or less conserved amino acids in the *E. coli* RF-2 region equivalent to domain IV of EF-G were mutated on the plasmid-borne copy of *E. coli prfB*. The mutant protein phenotypes were examined by using IPTG-controllable expression plasmid pSUIQ-RF2+ and its derivatives. Although most of these mutations simply reduced the activity of RF-2 for complementation and did not show any dominant phenotype over the wild-type protein, two mutations, *prfB2070* (Phe^{207} → Thr) and *prfB2130* (Arg^{213} → Ile), were toxic to wild-type cells when induced by IPTG (Table 1). These two positions are not conserved in prokaryotic and eukaryotic RFs, but their amino acids are conserved distinctively in each RF group (coordinates 269 and 275; positions marked by daggers in Fig. 2). *prfB2070* and *prfB2130* substituted the amino acids conserved in RF-1s for those conserved in RF-2s. Since expression of these mutations did not result in suppression of any nonsense alleles, we assumed that the dominant lethality may be caused by abnormal termination at sense codon(s) by the mutant RF-2. The amino acids in EF-G equivalent to these two substitutions are located at the end of β-strand 2 and the beginning of β-strand 3 in domain IV (positions marked by daggers in Fig. 3*A*). These positions are likely equivalent to the anticodon loop of tRNA (Fig. 4). Given that RFs form a similar topology whether or not lacking the counterpart moiety of β-strand 1 in the assigned gap (see Fig. 3*A*), these results suggest that *prfB2070* and *prfB2130* mutations affect the anticodon mimicry of the protein. Comparisons of multiple sequences for EF-Gs from distantly related species also revealed gap(s) at β-strand 1 (data not shown).

Although no published mutations have been mapped within the predicted tRNA-mimicry region of RF, four termination-altering substitutions, Gly^{180} → Ser (position 259), His^{182} → Tyr (position 261) of *S. typhimurium* RF-1, and Arg^{231} → Lys (position 265), Pro^{233} → Leu (position 267) of *S. cerevisiae* mRF-1, cluster at the predicted boundary between domains III–IV (see Fig. 2 and Table 1). One can assume that these mutations alter the topology of the tRNA-like rod structure by affecting the hinge linking the two domains. We suspect that the protein's putative anticodon determinant plays a direct role in stop codon recognition.

Ternary Complex Model for RFs. The overall resemblance of the EF-Tu–tRNA–GDPNP ternary complex and EF-G suggests common requirements for the structure and function on the ribosome. Of six subdomains of EF-G, C-terminal domains III–V are equivalent to RF-1 and RF-2 as to mimicking tRNA, whereas the three most N-terminal domains, G, G′, and II, that are equivalent to EF-Tu (34, 35), are not conserved in RF-1 and RF-2, but are highly conserved in bacterial RF-3 (34, 35) (see Fig. 4). Therefore, the overall protein moiety of EF-G seems to consist of the N-terminal RF-3 moiety and the C-terminal RF-1 or RF-2 moiety with some overlaps in domain III, suggesting the apparent resemblance of the RF-1/2–RF-3 complex, if formed appropriately, to EF-G. This raises the possibility that RF-3 may be the structural and functional homolog of EF-Tu, which brings tRNA-mimicking RF-1 and RF-2 to the A site of the ribosome by forming the ternary complex with RF-1/2 and GTP (ternary complex model).

Several lines of evidence favor this view. (*i*) RF-3 stimulates binding of RF-1 and RF-2 to the ribosomal termination complex by functionally interacting with these RFs and enhances peptide chain termination *in vivo* and *in vitro* (refs. 5,

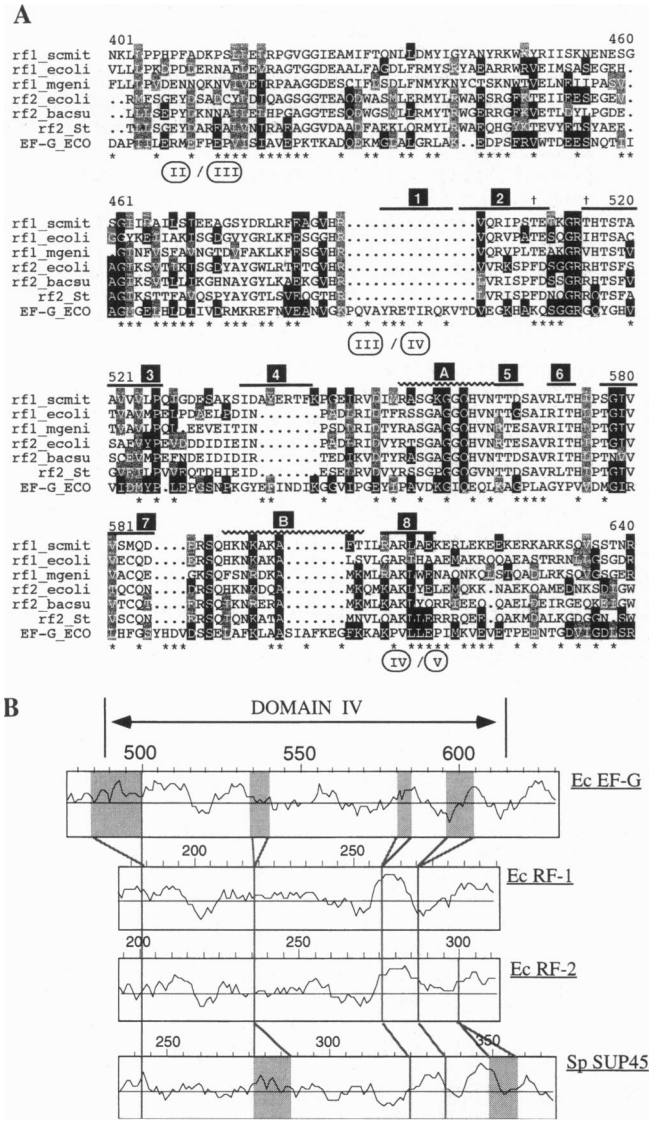

FIG. 3. (*A*) Comparison of the amino acid sequences of prokaryotic RFs and EF-G of *E. coli*. Similarity alignment was performed as described in Fig. 2, and identical or conserved amino acids were assigned. The number of the amino acid position is counted from the N-terminal Met of EF-G. Boundaries between structural domains II-V of EF-G, as revealed by x-ray crystallography, are marked under the sequence. The β strands and α helices in domain IV are indicated by bold and wavy lines, respectively, according to the reported assignment (34, 35). Daggers indicate positions of two dominant lethal mutations in *E. coli prfB* isolated in this study. Asterisks indicate amino acids conserved in at least five of six different EF-Gs: *Sulfolobus acidocaldarius* EF-2, *Thermoplasma acidophilum* EF-2, *E. coli* EF-G, *T. thermophilus* EF-G, soybean chloroplast EF-G, and rat mitochondrial EF-G. (*B*) Hydrophilicity profiles of predicted tRNA-mimicry regions, domain IV of EF-G and RF counterparts in *E. coli* RF-1, RF-2, and *Sch. pombe* Sup45. Horizontal numbers refer to amino acid residues, and shaded regions in EF-G and SUP45 correspond to gaps in the RF-1 and RF-2 sequence alignment. Areas above the median line are hydrophilic, and areas below are hydrophobic.

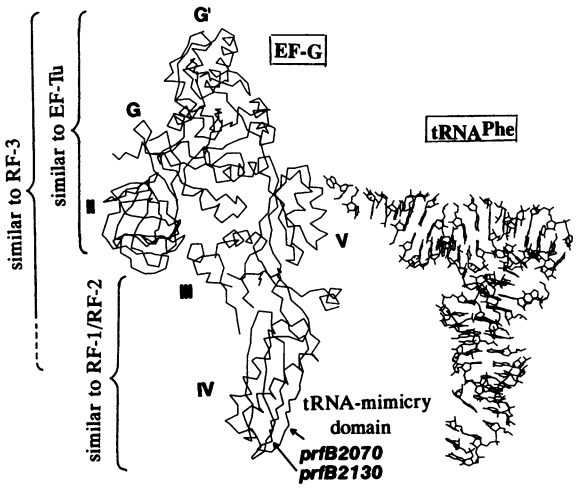

FIG. 4. Schematic diagrams of the structure of *T. thermophilus* EF-G (*Left*) and yeast $tRNA^{Phe}$ (*Right*) as well as predicted domain homology to bacterial RFs and EF-Tu. The backbone of EF-G is presented according to the C^{α} coordinates of the structure of EF-G in complex with GDP (provided by A. Liljas, University of Lund), and the structure of $tRNA^{Phe}$ is presented with the backbone and the side chains (provided by M. Tateno and S. Yokoyama, University of Tokyo). Both EF-G and $tRNA^{Phe}$ structures are drawn to the same scale. The G′ domain has been proposed to function as an intrinsic exchange factor modulating the binding of the guanine nucleotides and facilitating their exchange (34). Domain II is conserved among translation factors and together with the G domain makes a common structural unit possibly responsible for similar interactions with the ribosome (34). RF-1 and RF-2 show homology to domain IV and parts of domains III and V of EF-G; RF-3 shows homology to domains G, G′, II, III, and part of domain IV of EF-G; EF-Tu shows homology to domains G, G′, and II of EF-G. Amino acid positions equivalent to the dominant lethal *prfB* mutations are indicated.

6, and 41; K. Matsumura, K.I., and Y.N., unpublished data). (*ii*) The two RFs of *S. cerevisiae*, Sup45 (eRF-1) and Sup35 (eRF-3 candidate) also interact physically and functionally (12, 42). (*iii*) Guanine nucleotides switch RF-3 conformation between two alternative states: ribosome association (GTP-bound form) or ribosome dissociation (GDP-bound form) (J. Moffat, W. Tate, and Y.N., unpublished data). (*iv*) The activity of the eRF-1:eRF-3 complex for polypeptide release was also GTP-dependent (12). These observations support the ternary complex model of RFs. The predicted mimicry of bacterial RF-1/2–RF-3 complex to EF-G could be extended to eukaryotic translation factors, eRF-1, eRF-3, and a eukaryotic counterpart of EF-G, eEF-2.

Nevertheless, there are some apparent differences between bacterial RF-3 and yeast Sup35. (*i*) Overexpression of both Sup45 and Sup35 is required to enhance the efficiency of termination in *S. cerevisiae* (42), while overexpression of RF-1, RF-2, or RF-3 alone is sufficient to generate an antisuppressor phenotype in *E. coli* (41, 43). (*ii*) Yeast Sup35 is encoded by an essential gene while bacterial RF-3 is encoded by a nonessential gene. (*iii*) Yeast Sup35 and Sup45 bind *in vivo* (using the two-hybrid system) and *in vitro* (using immobilized Sup45 to precipitate Sup35) and exist as a heterodimer in yeast cell lysates (42), while bacterial RF-3 does not bind appreciably to RF-1 or RF-2 using the same experimental procedures (Y. Kawazu, K.I., and Y.N., unpublished data). (*iv*) Sup35 shows considerable C-terminal homology to EF-1α (for a review, see ref. 11), while RF-3 has a long C-terminal polypeptide compared with EF-Tu, which shows significant homology to domain III (41) and part of domain IV (ref. 41; K.I. & Y.N., unpublished data) of EF-G. Therefore, it cannot be excluded at present that bacterial RF-3 is a RF-1/RF-2-specific EF-G homolog and stimulates translocation of a tRNA-like protein in contrast to Sup35 which perhaps acts as an EF-1α homolog to form the ternary complex. Finally, Sup35 is a yeast cytoplasmic prion-like element called [*psi*$^+$] (44, 45). The [*psi*$^+$] determinant was uncovered some 30 years ago as a modifier of tRNA-mediated nonsense suppression in *S. cerevisiae* that was inherited in a non-Mendelian fashion (46). Sup35 has several tandem peptide repeats with the consensus to other prion proteins of mammals (for a review, see ref. 11). Therefore,

Sup35 is likely to be able to assume two functionally distinct conformations that differentially influence the efficiency of translation termination. Further definition of the roles of bacterial RF-3 and yeast Sup35 will provide us with clues to the mechanisms involved in prokaryotic and eukaryotic polypeptide termination.

The model of a tRNA–mimicry domain in the RF protein structure could explain how RFs have the ability to recognize the stop codon and the A site of the ribosome. Further analysis of the role(s) of conserved amino acids will provide us with clues to protein–nucleic acid mimicry and the mechanism underlying stop codon recognition, which has been a long-standing coding problem.

We are grateful to A. Liljas for the C^{α} coordinates of the structure of EF-G in complex with GDP (2.4-Å resolution), to M. Tateno and S. Yokoyama for expertise in computer graphics, and to A. Ogiwara for the discussion of possible sequence alignments. We thank M. Tuite for the gift of *S. cerevisiae SUP45* clones and mutants, C. Olsson and D. Court for critical reading of the manuscript, and C. Yanofsky for many valuable comments and criticisms on the manuscript. This work was supported in part by grants from the Ministry of Education, Science and Culture, Japan, and the Human Frontier Science Program (awarded in 1993).

1. Scolnick, E., Tompkins, R., Caskey, T. & Nirenberg, M. (1968) *Proc. Natl. Acad. Sci. USA* **61,** 768–774.
2. Capecchi, M. R. & Klein, H. A. (1969) *Cold Spring Harbor Symp. Quant. Biol.* **34,** 469–477.
3. Caskey, T., Scolnick, E., Tompkins, R., Goldstein, J. & Milman, G. (1969) *Cold Spring Harbor Symp. Quant. Biol.* **34,** 479–488.
4. Goldstein, J. L. & Caskey, C. T. (1970) *Proc. Natl. Acad. Sci. USA* **67,** 537–543.
5. Mikuni, O., Ito, K., Moffat, J., Matsumura, K., McCaughan, K., Nobukuni, T., Tate, W. & Nakamura, Y. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5798–5802.
6. Grentzmann, G., Brechemier-Baey, D., Heurgue, V., Mora, L. & Buckingham, R. H. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5848–5852.
7. Craigen, W. J., Lee, C. C. & Caskey, C. T. (1990) *Mol. Microbiol.* **4,** 861–865.
8. Tate, W. P. & Brown, C. M. (1992) *Biochemistry* **31,** 2443–2450.
9. Konecki, D. S., Aune, K. C., Tate, W. & Caskey, C. T. (1977) *J. Biol. Chem.* **252,** 4514–4520.
10. Frolova, L., Le Goff, X., Rasmussen, H. H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A.-L., Celis, J. E., Philippe, M., Justesen, J. & Kisselev, L. (1994) *Nature (London)* **372,** 701–703.
11. Stansfield, I. & Tuite, M. (1994) *Curr. Genet.* **25,** 385–395.
12. Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L. & Philippe, M. (1995) *EMBO J.* **14,** 4065–4072.
13. Stansfield, I., Grant, C. M., Akhmaloka & Tuite, M. F. (1992) *Mol. Microbiol.* **6,** 3469–3478.
14. Sikorski, R. S. & Hieter, P. (1989) *Genetics* **122,** 19–27.
15. Ogawara, H., Urabe, H., Ohtaki, R. & Nakamura, Y. (1995) *J. Bacteriol.* **177,** 5342–5345.
16. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 488–492.
17. Tuite, M., Akhmaloka, Firoozan, M., Duarte, J. A. B. & Grant, C. M. (1991) in *Post-transcriptional Control of Gene Expression*, eds. McCarthy, J. & Tuite, M. (Springer, Berlin), pp. 611–622.
18. Craigen, W. J., Cook, R. G., Tate, W. P. & Caskey, C. T. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 3616–3620.
19. Elliott, T. (1989) *J. Bacteriol.* **171,** 3948–3960.
20. Kawakami, K., Jönsson, Y. H., Björk, G. R., Ikeda, H. & Nakamura, Y. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 5620–5624.
21. Kawakami, K. & Nakamura, Y. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 8432–8436.
22. Fraser, C. M., Gocayne, J. D., White. O., Adams, M. D., *et al.* (1995) *Science* **270,** 397–403.
23. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., *et al.* (1995) *Science* **269,** 496–512.
24. Pel, H. J., Rep, M. & Grivell, L. A. (1992) *Nucleic Acids Res.* **20,** 4423–4428.
25. Devereux, J., Haeberili, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
26. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington), Vol. 5, pp. 345–352.
27. Moffat, J. G., Timms, K. M., Trotman, C. N. A. & Tate, W. P. (1991) *Biochimie* **73,** 1113–1120.
28. Mikuni, O., Kawakami, K. & Nakamura, Y. (1991) *Biochimie* **73,** 1509–1516.
29. Mironova, L. N., Samsonova, M. G., Zhouravleva, G. A., Kulikov, V. N. & Soom, M. J. (1995) *Curr. Genet.* **27,** 195–200.
30. Elliott, T. & Wang, X. (1991) *J. Bacteriol.* **173,** 4144–4154.
31. Breining, P. & Piepersberg, W. (1986) *Nucleic Acids Res.* **14,** 5187–5197.
32. Pel, H. J., Rep, M., Dubbink, H. J. & Grivell, L. A. (1993) *Nucleic Acids Res.* **21,** 5308–5315.
33. Zhang, S., Rydén-Aulin, M., Kirsebom, L. A. & Isaksson, L. A. (1994) *J. Mol. Biol.* **242,** 614–618.
34. Ævarsson, A., Brazhnikov, E., Garber, M., Zheltonosova, J., Chirgadze, Yu., Al-Karadaghi, S., Svensson, L. A. & Liljas, A. (1994) *EMBO J.* **13,** 3669–3677.
35. Czworkowski, J., Wang, J., Steitz, T. A. & Moore, P. B. (1994) *EMBO J.* **13,** 3661–3668.
36. Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B. F. C. & Nyborg, J. (1995) *Science* **270,** 1464–1472.
37. Doudna, J. A., Cech, T. R. & Sullenger, B. A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 2355–2359.
38. Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., Mosbaigh, D. W. & Tainer, J. A. (1995) *Cell* **82,** 701–708.
39. Savva, R. & Pear, L. H. (1995) *Nature (London) Struct. Biol.* **2,** 752–757.
40. Liljas, A. (1990) in *The Ribosome: Structure, Function and Evolution*, eds. Hill, W. E., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, D. & Warner, J. R. (Am. Soc. for Microbiol., Washington), pp. 309–317.
41. Kawazu, Y., Ito, K., Matsumura, K. & Nakamura, Y. (1995) *J. Bacteriol.* **177,** 5547–5553.
42. Stansfield, I., Jones, K. M., Kushnirov, V. V., Dagkesamanskaya, A. R., Poznyakovski, A. I., Paushkin, S. V., Nierras, C. R., Cox, B. S., Ter-Avanesyan, M. D. & Tuite, M. F. (1995) *EMBO J.* **14,** 4365–4373.
43. Weiss, R. B., Murphy, J. P. & Gallant, J. A. (1984) *J. Bacteriol.* **158,** 362–364.
44. Doel, S. M., McCready, S. J., Nierras, C. R. & Cox, B. S. (1994) *Genetics* **137,** 659–670.
45. Ter-Avanesyan, M. D., Dagkesamanskaya, A. R., Kushnirov, V. V. & Smirnov, V. N. (1994) *Genetics* **137,** 671–676.
46. Cox, B. S. (1965) *Heredity* **20,** 121–521.